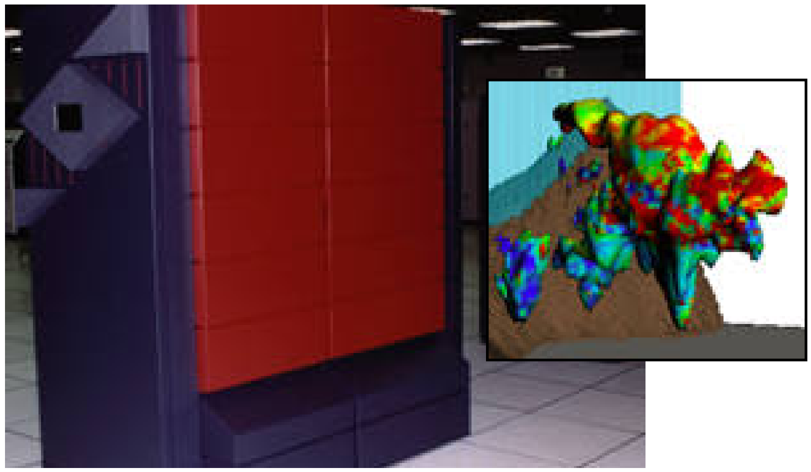
The Industrial Computing Initiative



Our three-year, multiparty collaboration is addressing several different problems that have limited more widespread use of massively parallel computing by researchers in government, academia, and industry. The delivery of a set of tools and efficient applications that can be run on different machines will accelerate the use of high-performance parallel processing to increase U.S. industrial productivity and competitiveness.

FOR more than four decades, the strength of LLNL has been based on large-scale facilities and associated science teams working to make bold advances in science and technology. One of the areas in which we have become world renowned for our resident expertise is high-performance computing.

The history of LLNL intersects strongly with the history of computing. We have made important contributions in software, operating systems, scientific applications, and computing techniques. We also have a history of successful partnerships with private industry and other government laboratories.

Today, many computational science projects within the DOE laboratories and U.S. industry are facing a challenge. To move forward, the projects need to add further realism, which will come from augmenting physical effects, resolution, or dimensionality. Increasing realism will expand the demands on a computational resource by orders of magnitude. This demand has driven the movement toward the use of computers with multiple processors. These processors, working together, can rapidly solve a single problem. This concept is called parallel processing.

However, massively parallel computing has not been adopted as the high-end standard in computational research as rapidly as it might have because of the difficulty in creating efficient, high-performance parallel programs. The obstacles do not arise from a single source. Rather, they are due to deficiencies both in the hardware design and in the software programming environment of virtually all massively parallel systems. In addition, each vendor offers a unique architecture and creates unique software products. Such variations, along with other issues described in this article, now

Glossary

Term	Definition
Central processing unit (CPU)	The part of the computer containing the circuits required to interpret and execute the instructions.
CMOS	Complementary metal oxide semiconductor.
Flops	Floating-point (arithmetic) operations per second; a commonly used measure of the speed of calculation.
H4P	High-Performance Parallel Processing Project.
ICI	Industrial Computing Initiative.
Latency	The waiting time between the issuing of a read instruction and the receipt of requested data.
Massively parallel processor	A parallel processing machine with 100 or more microprocessor-based CPUs.
Microprocessor	A single silicon chip on which the arithmetic and logic functions of a computer are placed. A typical microprocessor contained about 35,000 devices in 1982; recent ones contain about 3.5 million.
Node	An intersection point in the communication topology of a massively parallel processor. The CRAY T3D has two processors per node, and the communication topology is a three-dimensional torus.
Parallel processor	A machine that uses more than one processor running simultaneously to speed up the solution of a computational problem.
Pipelining	The computer architect's version of an assembly line. Instructions are overlapped in execution, and a new operation is started every clock cycle even though it takes several clock cycles to complete one operation.
Porting	Moving an application code from one computer system so that it runs on another of a different type, for example, from the Thinking Machines CM-5 to the CRAY T3D or vice versa.
Production environment	All of the system components needed for a user to develop an application efficiently, debug it, execute it, and assimilate the output. These components include the computer itself along with one or more graphics workstations, high-speed networks, high-performance storage systems, and documentation, consulting, and software tools.
Scalar code	An application that is not, or cannot be, vectorized is called scalar. (See vector processor.)
Supercomputer	A computer that is among those with the highest speed and largest memory at any given time.
Vector processor	A computer with hardware instructions that can each operate on a set of data elements, achieving high speed by streaming the set of elements through the hardware segments in a pipelined fashion. (See pipelining.) A code executing in this mode is called a "vectorized" code.
VLSI	Very large-scale integration.

impede the movement of parallel applications across different platforms. Investment in the new technology both by government laboratories and private industry will accelerate as the difficulties are overcome.

The Industrial Computing Initiative (ICI) represents a collaborative effort by major industrial partners and government laboratories to develop applications targeted at parallel computers. Leverage can come from working in common; in particular, the ICI effort is of sufficient scale to allow for a general and more complete assessment of massively parallel technology.

The ICI is one part of a broader project called the High-Performance Parallel Processing Project (H4P). The H4P, valued at $66 million over three years, is funded by the DOE and private industry in a 50-50 cost-shared manner. The ICI portion of the project, in which the Laboratory is playing a key role, is valued at $52 million.

What Has Changed?

The current situation in high-performance computing has been compared to what troubled the railroad industry in the 19th century when different gauges of track prevented the transport of goods on different lines. Now, industry is experiencing the computational equivalent of that situation. To understand how we arrived at this place, some basic definitions and a brief look at the unprecedented growth of computing power over the last few decades will be helpful.

When the first commercial computers came on the market in the 1950s, each user had access to an entire machine—its processor, memory, and storage. In a sense, a user temporarily "owned" a machine while others waited their turn. By the mid 1960s, the idea of timesharing came into being. In this approach, the operating system directs the central processing unit (CPU) to work on several jobs in tandem. Instead of letting the CPU sit idle while a time-consuming step in one job is completed (often the input/output function), the operating system juggles jobs in and out of the CPU to make the most efficient use of a single processor and to give each user a sense of interactive control.

Modern conventional supercomputers provide not only time-shared access but also extremely rapid computational performance. One measure of performance is speed of calculation, usually expressed as Mflops, which is a million floating-point (or arithmetic) operations per second. Most current supercomputers achieve hundreds or sometimes thousands of Mflops by streaming a set of data elements through the hardware segments in a pipelined fashion.

Parallel computing represents a leap forward in the efficient and potentially flexible use of processor resources to solve computational problems. Parallel computing is simply the simultaneous execution of operations, transmission of information, or storage of data through the use of multiple processors. In massively parallel processing (MPP), large numbers of processors are used to attack an even larger set of tasks that together compose the problem to be solved. Whereas current conventional supercomputers still use one or a few powerful CPUs, massively parallel machines can have 100, 1000, or even more microprocessor-based CPUs. There is no fixed boundary between parallelism and massive parallelism, but a machine with more than 100 CPUs is generally called massively parallel.

The transition to massively parallel systems has been driven by the performance revolution in microprocessor technology. When personal computers came on the market some 15 years ago, they sparked a revolution. The "brain" in a personal computer consists of a simple computing device on a single silicon chip, the microprocessor. Early microprocessors were slow by current standards and were useful only for limited applications.

In the past decade, microprocessor technology has improved rapidly in both absolute and relative performance. One measure of absolute performance is the clock rate—the rate at which a processor operates, also called the processor cycle time. For microprocessors, the clock rate improved by a factor of about 50 from 1982 to the present. In contrast, the clock rate for conventional supercomputers improved over the same period by a factor of about 3. Microprocessor clock rates will exceed those of supercomputers within a couple of years.

In terms of relative performance, the story is similar. Figure 1 shows the evolution of performance for microprocessors versus conventional supercomputers.[1] In this graph, the measure of performance is the peak rate for floating-point operations on a common benchmark problem. The two curves show that microprocessors are closing the performance gap with supercomputers and will eventually take the lead. The comparison becomes increasingly sobering for traditionalists when the measure used for comparison is based on cost performance.

In a recent study conducted at LLNL, the performance of computers over the years was measured in terms of dollars per mips (millions of instructions per second).[2] Depending on the type of technology, it takes from one to a few instructions to complete a floating-point operation. Figure 2 shows that a sharp change in the slope of the cost-effectiveness curve occurred in the late 1980s.

We can conclude from such data that parallel machines, leveraging off commodity processors, have an inherent advantage in cost performance. Moreover, distributed-memory MPP machines using microprocessor technology have three other built-in advantages over current vector supercomputers. They offer higher peak speeds (for systems with larger numbers of processors), larger aggregate memory, and, in some cases, abundant software designed for the workstation market. Putting all this information together, the potential advantage for parallel computers appears to be overwhelming. For example, an examination of results from a widely accepted benchmark suite (the NAS Parallel Benchmarks[4]) shows that the latest parallel machines are superior for several, but not all, types of problems. However, the dominance of parallel machines has been slower to emerge and is not yet as overwhelming and complete as might have been expected from optimistic forecasts.

What Is the Problem?

We have seen that massively parallel machines are becoming increasingly cost effective relative to established supercomputers. Moreover, MPP systems can motivate larger, more sophisticated codes, in part, because they support a much larger memory than a conventional vector supercomputer. Why then isn't MPP technology already in wide use in the industrial sector? In a nutshell, the problem is a combination of unresolved technical challenges in both the hardware and software.

The most serious problem facing parallel computers is that each processor has a local memory. When one processor requires information available on another processor's local memory, a communication process must be effected. This represents an extra dimension of complexity—one that is built into a parallel machine—and is largely absent from a sequential monoprocessor design. To complicate matters, the computation and communication hardware components for parallel processing are not designed to cooperate from the outset but, instead, are engineered to cooperate after the microprocessor has been designed (for the workstation). Cooperation between separately designed hardware components makes communication less efficient than would otherwise be the case. This situation is unavoidable because microprocessors are commodities.

During a calculation, every problem is broken down into tasks or fragments. Unless all tasks represent completely independent problems, the processors assigned to a specific task will, at some point, need to have information computed elsewhere. This need for information raises issues concerning communication between the processors, its rate and its latency (how fast does the first word get back?). MPP programmers worry about synchronization between tasks, the layout of data across the separate memories of all the processors, data transfer rates, and processor idle time. In other words, it is easy to design a parallel program that is inefficient either through lack of insight or experience on the part of a programmer or through poorly designed, developed, and tested software and hardware in the target system. Unless a balance exists among processor speed, latencies, communication rates, and input/output capabilities, a system will lack generality. In this case, only a small class of problems (those that

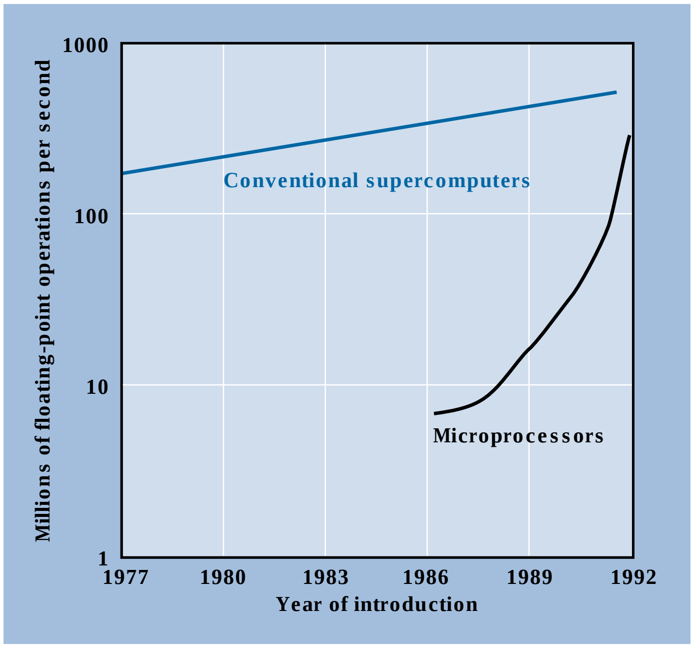


Figure 1. The evolution of performance for microprocessors vs conventional supercomputers central processing units.[1] The supercomputer curve shows a steady, but gradual, increase in performance over the last 15 years. Dramatic improvements in integrated circuit technology are allowing microprocessors to close the performance gap with conventional supercomputers.

do not tickle the system's weak points) runs well on that system.

Unfortunately, as the number of processors in a system increases, the number of tasks must increase as well, or some processors will be idled. With the increase, weaknesses in system design are aggravated and become more obvious, leading to decreased efficiencies, saturation, and, ultimately, defeat. The massively parallel processor lives in the niche of high processor counts, so it is the most vulnerable architecture with respect to the issue of balance.

Because each system must contend with a complicated architecture involving multiple processors, communication hardware, interconnect topology, and input/output processors, a vendor faces a myriad of design choices. As a result, each manufacturer today offers a system that is very different from any other competing system. The divergence of products begins at the hardware level. After layers of software, operating systems, programming models, debuggers, and performance-monitoring tools are added, the result is that codes designed for one system (even if ported with only a few weeks' work) will generally run poorly on another system until they are tuned. After all, the codes were not designed with the idiosyncrasies or weaknesses of some alternative system in mind.

To recap and add to the list of problems, the limitations and perceived weaknesses of MPP machines emanate from:

- Difficulty in programming efficiently and portably, especially for certain applications requiring extensive communication and abundant synchronization.
- The lack of sophisticated and time-tested support tools, such as debuggers and performance-analysis tools.
- Inadequate timesharing support, making it difficult for many users to share a machine efficiently for code development and short debugging runs.
- The lack of refined parallel physics, engineering, and chemistry application codes as well as efficient parallel mathematical libraries.
- The lack of efficient, standardized, and error-proof parallel input/output protocols and capabilities.
- A lag in processor capability. The latest microprocessor frequently

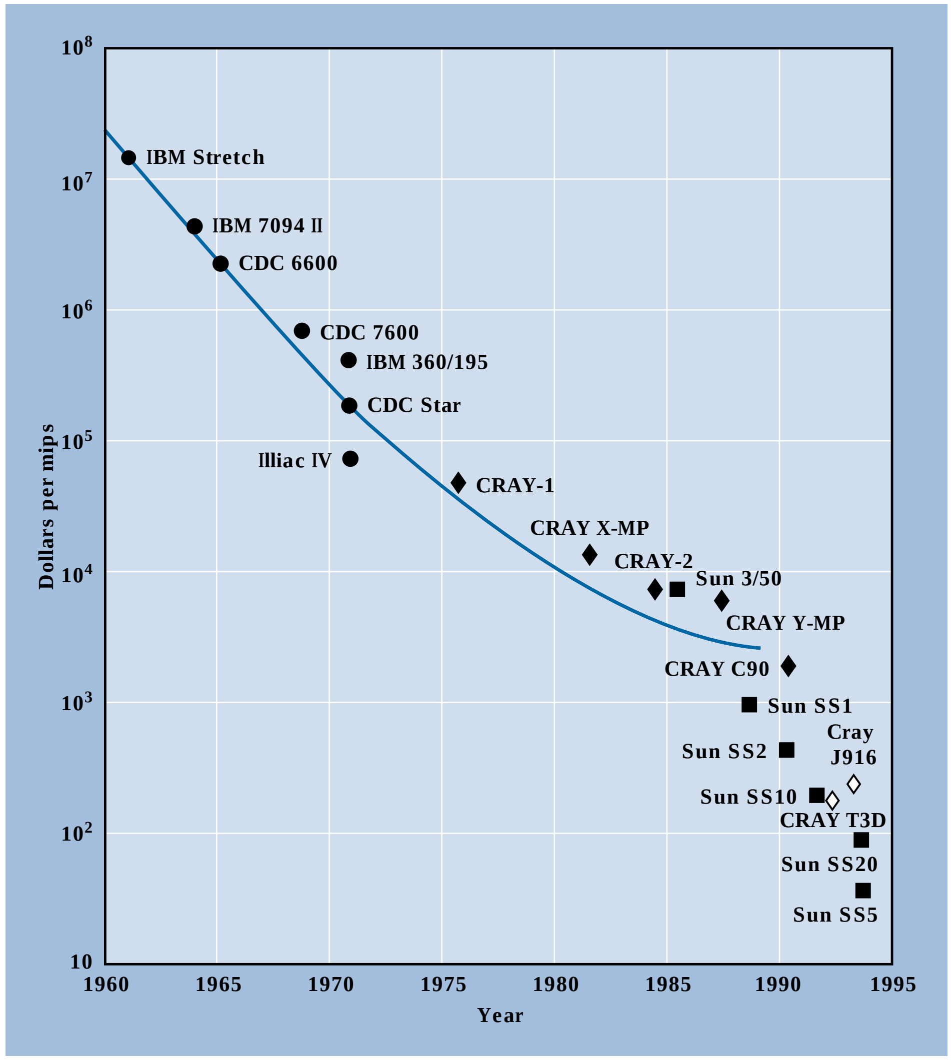


Figure 2. The performance of computers over the years measured in terms of dollars per millions of instructions per second (mips).[2] A sharp change in the slope of the cost-effectiveness curve occurred in the late 1980s. Solid circles are dollars per mips for conventional supercomputers from a study by R. Turn.[3] Solid diamonds are dollars per peak Mflops vs arrival time for supercomputers. (In this case, instructions per second is roughly equivalent to flops per second.) Solid boxes are the microprocessor-powered equipment from Sun Microsystems plotted as dollars per VAX mips as a function of year of availability. Open diamonds are the CMOS VLSI-powered equipment from Cray Research. The T3D is the CRAY distributed-memory massively parallel processor (using a single chip microprocessor) that is being used as part of the current ICI project. The CRAY J916 is a recent, shared-memory, CMOS-technology parallel processor.

does not reside in the latest MPP system because of the design time of the MPP itself.

Despite many years of development in parallel processing, major problems and challenges have effectively separated parallel processing research from the mainstream of supercomputing applications. Industry sees a gulf between what is required in real-world computing and the idealizations that have dominated the field of parallel computing research. Stated in simple terms, those companies that contemplate a move to massively parallel computing face a real dilemma: the risk of pursuing MPP, which entails a significant investment of resources, and the risk of falling behind international competition by not pursuing the most advanced computer technology.

U.S. industries simply do not have efficient applications, and they lack the necessary software tools and languages to facilitate code development in an MPP environment. Developing efficient applications that can realistically simulate complex problems and that have the potential to run on MPP hardware from different vendors is what the ICI is all about.

Figure 3. **The CRAY T3D after its recent installation at LLNL. This massively parallel computer includes a high-performance, 88-Gbyte disk subsystem. The memory of the T3D is logically shared, but physically distributed, within a three-dimensional toroidal array of nodes. All system memory is directly accessible to all 128 processors. Each microprocessor provides 150 Mflops of peak performance. The number of processors and the disk capacity will be doubled in December for the benefit of LLNL and University of California researchers.**

The Partners

The issues being addressed by the ICI demand an unprecedented level of innovation and cooperation from computer manufacturers, computer scientists, engineers, programmers, and participants from the industrial sector. Overall, the project will tap and put to use the expertise of more than 40 scientists from two national laboratories, six specialists in parallelization from Cray Research, two specialists from Thinking Machines Corporation, and at least 17 industrial scientists.

National Laboratories

The two national laboratories, Lawrence Livermore and Los Alamos, bring to the endeavor their expertise in supercomputing applications and MPP enhancements gained from large-scale simulation and modeling of complex phenomena. Through their weapons and supporting science programs, they have developed a unique science base that is necessary to solve large problems. The laboratories also offer extraordinary infrastructure support in both intellectual and physical terms. Thus, the two national laboratories are well positioned to fulfill the project's objectives and the DOE's mandate to transfer new technologies to the private sector.

The importance of computing at these facilities is revealed in the financial commitment to it. At LLNL, for example, some 10% of the nearly billion dollar annual budget is spent on the development of applications and system software, and nearly 10% of all LLNL employees work for the Computation Organization.

Cray Research

Cray Research has played a major role in the supercomputer industry for almost two decades and has a well-established collaborative relationship with the national laboratories. The company has made available for the project two T3D high-performance massively parallel computer systems (Figure 3). One of the two T3Ds is now sited at Livermore's National Energy Research Supercomputer Center (NERSC); the other is sited at Los Alamos' Advanced Computing Laboratory (ACL), a High-Performance Computing Research Center. These machines will eventually be linked together in a prototype distributed MPP computing environment.

The T3D hardware design permits a global address space implemented in the hardware. This design means that it is possible for a programmer to view the memory as shared, not distributed. A programmer must still worry about indiscriminate addressing of memory because an off-processor fetch still takes about six times as long as a local fetch. However, the machine's designers focused on the requirements of balance, and the machine's interprocessor communications network today defines the state of the art.

The T3D hardware sets the stage for significant software improvements. In particular, it permits efficient implementation of all three programming models that are now commonly used (the models are called message passing, data parallel, and work sharing). This feature is important because all three programming models can be implemented in one application. Thus, scientists can choose an appropriate combination of models that best fits the algorithms employed, rather than trying to fit the application to a single model provided by a hardware vendor.

Thinking Machines

Thinking Machines Corporation is a small, privately held company that established itself as a major MPP system supplier over the last 5 years. This company was the first to demonstrate the routine use of tens of thousands of processors on a single program. It was also the first to demonstrate a scalable programming standard for parallel computing. Its advanced massively parallel system, the 1024-node Connection Machine (CM-5) supercomputer system, is currently at the Los Alamos ACL and will be made available to the project team.

Industrial Partners

The industrial partners are major U.S. companies spanning a spectrum of commercial enterprises. Among the partners with whom LLNL is negotiating CRADAs are: AT&T, Alcoa, Boeing, Hughes, Halliburton, Areté, IT Corporation, and Xerox. The issues these companies need to address are ones of national concern. They range from protecting and preserving the environment to developing advanced materials, improving manufacturing techniques, and enhancing the nation's high-performance computing infrastructure. Some 17 scientists from industry will participate in the ICI.

Project Objectives

The central objective of the three-year project is to accelerate the development of parallel computing so that it can serve as an effective competitive tool for U.S. industry. This development will be implemented by a selected set of industrial applications that will have practical use and that are written in portable programming languages.

Whereas the effort addresses many different technical issues, the principal objectives fall into two broad areas:

- Critical applications development to address problems that are of interest to the laboratories and that industry has identified as top priority.
- Accelerated development of integrated production MPP environments, including the development of portability tools and standards that will allow applications to run on hardware from different vendors.

Critical Applications

We have chosen to demonstrate the range and depth of science that is accessible through MPP by developing a few key, representative applications. Some of the highly complex applications that fit this criterion are, for example, modeling pollution transport in the environment (Figure 4), tools to study materials around boreholes for petroleum exploration (Figure 5), advanced materials design (Figures 6 and 7), and advanced manufacturing problems (Figures 8 and 9).

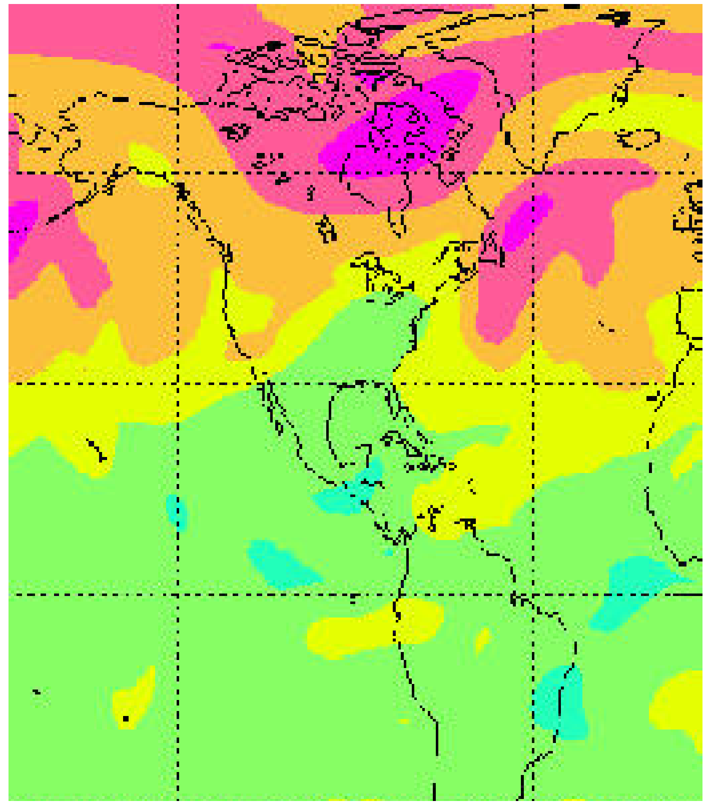
Many of the applications we are addressing had their origins in defense programs. They will now serve a dual use in that enhanced versions of the codes can be moved back into defense work and serve as industrially useful software products as well. The box on pp. 10–11 summarizes the various applications on which we will focus at first. As the collaboration develops, work should expand to include a wider range of applications.

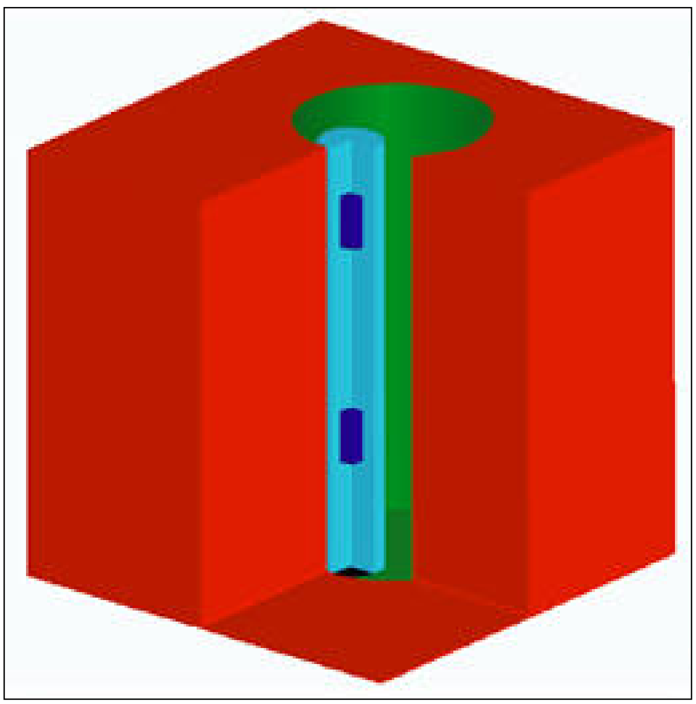
Complex applications codes usually consist of various modules, each of which treats a different aspect of the physics or other work, such as the numerical algorithm used to advance equations towards solution. We now know that each module or algorithm of an application maps best to a particular programming model. Nonetheless, very few applications programs actually use multiple programming models; efficiency suffers as a result of having to adhere to a single programming model because of limitations in hardware or in the software environment. As mentioned earlier, one of the many attributes of the newer machines is an ability to use two or three different programming models (such as message passing, data parallel, or work sharing) in one application. Thus, another aspect of developing efficient applications will be to make efficient use of available programming models.

Integrated Production Environments

A massively parallel computational environment is made up of many components, including graphics workstations, high-speed networks, high-performance storage systems, and software tools, as well as the supercomputers themselves. If a user can efficiently develop an application, debug it, execute it, and assimilate the output, a production environment exists. If a user cannot,

Figure 4. In one ICI effort, LLNL researchers are developing a global atmospheric chemistry model. Highly complex problems, such as environmental modeling, can be best addressed by the new MPP machines now arriving in the marketplace. This work can help guide industry in developing and analyzing products through a better understanding of their possible environmental consequences. The simulation shown here is a model of the distribution of ozone during the month of December at an altitude of about 20 km above Earth's surface. ▶

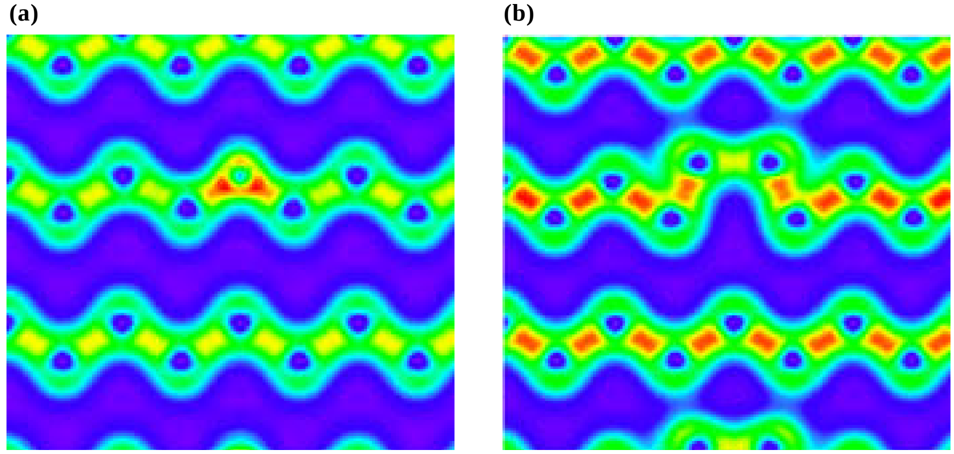




◀ **Figure 5.** We are developing MPP computer codes to calculate the response of nuclear logging tools used to study material surrounding exploratory boreholes in the quest for petroleum. This simulation shows the typical configuration of a borehole logging tool. Neutrons from a point source (black dot) at the bottom of the tool (light blue) are scattered in the borehole (green) and surrounding formation (red) before striking detectors (dark blue) on the tool.

Figure 6. Efficient materials modeling tools in MPP environments will benefit many sectors of the industrial community. Such tools can help to accelerate the synthesis–processing–fabrication–manufacturing cycle and ultimately contribute to the goal of materials by computer design. These illustrations show how electrons (red) interact with (a) defects and (b) impurities in silicon devices during the processing step. By understanding such interactions, we can better control processing temperatures and procedures to increase the efficiency of manufactured silicon devices. ▶

(a) (b)



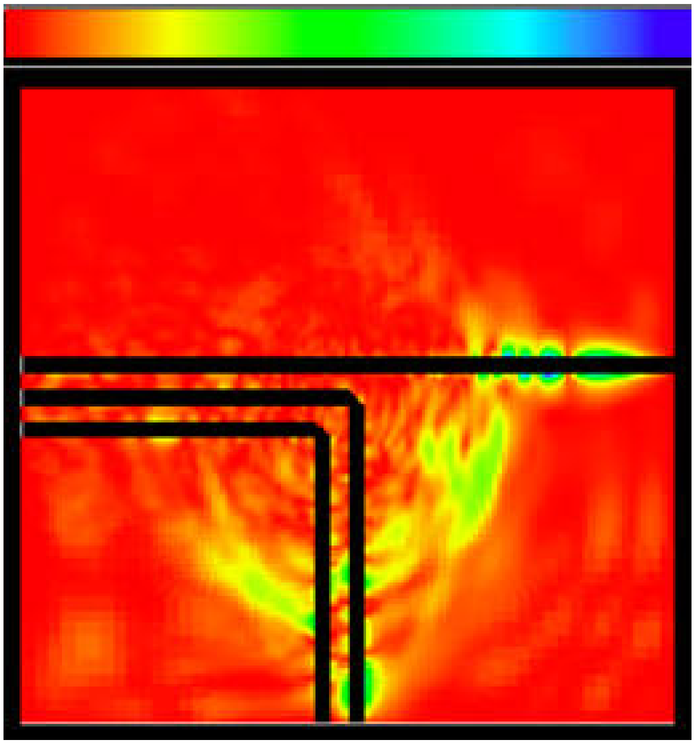


Figure 7. Another ICI project involves microwave components design. In this simulation, the color scale at the top represents the intensity of an electric signal, ranging from weak (red) to very strong (blue). The square below the scale represents a circuit board containing a three-way interconnect. Each microstrip conductor (black lines) is a few tenths of a millimeters in size. A test signal with a pulse width of two trillionths of a second should ideally propagate from left to right along the center conductor and follow the 90-degree corner without dispersion to the adjacent conductors. Instead, the electric fields show cross-coupling to the other conductors, particularly evident on the upper conductor (notice blue on the right-hand side). This type of simulation illustrates the difficulties that can arise in moving to higher-frequency signals that are needed for faster operation of devices.

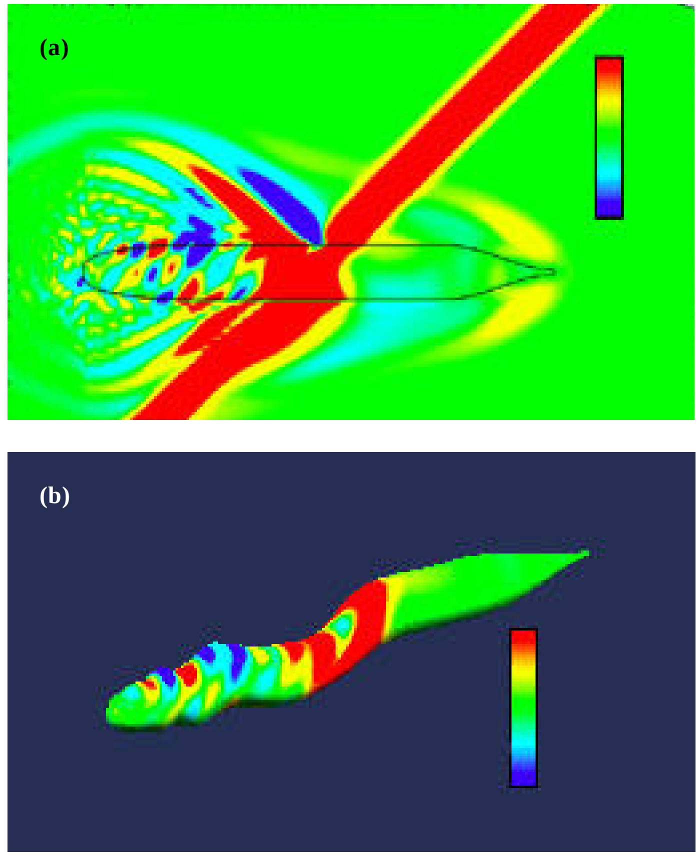


Figure 8. Our work in fluid–structure interactions will create a new simulation capability in the area of structural acoustics by the shipbuilding and aerospace industries. (a) In this snapshot of total pressure, an incident pressure wave (red) is midspan on a simulated elastic body (black outline). The pressure field highlights the presence of both a supersonic structural wave and flexural waves on the elastic body. (b) When the total pressure is mapped onto the surface of the elastic body itself, the exaggerated displacements help to show the presence of the flexural waves on the body.

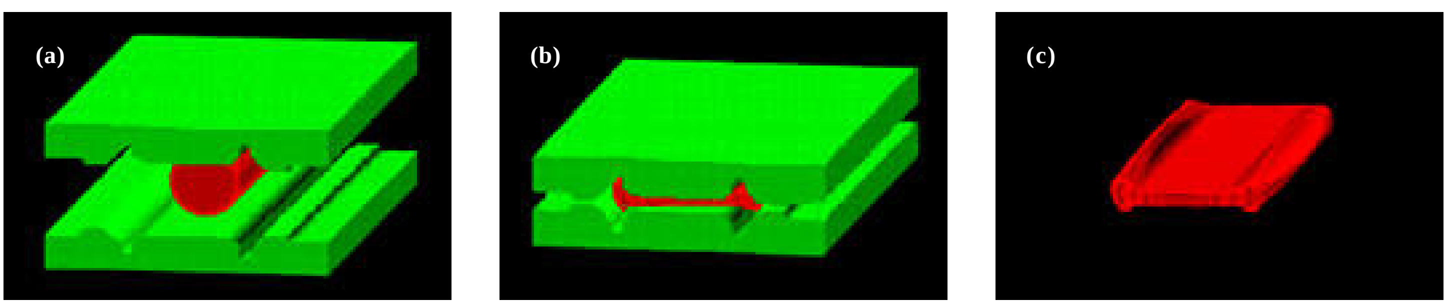


Figure 9. MPP codes for fluid dynamics and structural response will allow the U.S. metals industry to develop improved forging, extrusion, and casting processes. Here a three-dimensional code treats complex flows for a variety of materials. (a) At the start of a forging process, an aluminum rod (red) is positioned in a die (green). The forging process is represented in (b), and the final result is (c) an aluminum billet. Although the simulation represented here was run on a CRAY Y-MP machine, the code is currently running on the new CRAY T3D supercomputer at LLNL.

such an environment does not exist. At present, a production environment for parallel computers does not exist. As sophisticated as they may be, many industrial and DOE scientists do not enjoy working in a developing computational environment.

The ICI project will include a comprehensive effort to build an enhanced infrastructure for large-scale computing. The activities include developing high-speed network connections, a new multi-user scheduling system, code portability standards and techniques, performance enhancement tools, and visualization tools. A collaboration of systems and application developers will ensure that the final product is a fully tested MPP environment that will support the development of applications as well as large-scale production.

First, a mechanism is needed that allows many programmers to work on a machine simultaneously and to develop code efficiently. One answer is a space- and time-shared scheduler. Without a scheduler, it is impossible to halt a long-running program to allow others to use the processors without cancelling the original job. It is impossible to use machine partitions efficiently for code development because no mechanism is available to share the processors.

The DOE national laboratories have developed scheduler software for earlier massively parallel

What We Can Expect from the New MPP Applications

The Industrial Computing Initiative will initially focus on problems that can be best addressed and solved by the low-latency, tightly coupled, large-memory computer systems now arriving on the marketplace. Whereas the applications address a very broad range of topics, they share a common objective: helping the U.S. to compete more effectively in the global marketplace. All of the following topics are of interest to the DOE, the national laboratories, and the manufacturers of MPP machines as well. The benefits associated with each project are explained in the right-hand column. The nine projects involving LLNL are identified as blue type, and each one is described in this article.

Environmental Modeling

Lithography characterization for remediating underground pollution	Will reduce the costs of characterizing and remediating sites with underground contamination.
Unstructured grids for three-dimensional (3D) representation of heterogeneous materials	Will help industry cost-effectively design internal combustion engines, contain pollution, remediate groundwater contamination, and treat contaminated surface water.
Subsurface flow and chemical migration	Will allow industry to characterize the migration and transformation of contaminants in soils at waste sites. A new 3D simulation capability will improve the design and management of engineered remediation strategies.
Global atmospheric chemistry models	Will guide industry in developing and analyzing products through a better understanding of their possible environmental consequences.

Petroleum Applications

General reservoir simulation	Will allow oil exploration and oil service companies to perform simulations at fine grid resolutions.
Nuclear well logging	Will provide computer codes to calculate the response of nuclear logging tools used to study material surrounding exploratory boreholes.

machines—in particular, for the BBN TC-2000 as part of the LLNL Massively Parallel Computing Initiative. A similar, general scheduled environment is planned for the T3D, as shown in Figure 10. Developing a scheduler rapid enough for interactive code development will be complicated by the inability of the T3D processors to context switch (this means that multiple processes cannot remain resident in memory). Therefore, an entire job must be swapped to another system (called the front end) or to disk to allow access to the processors by waiting processes. Although swapping is possible, it is also more time consuming than context switching, so the resulting environment will probably be less than optimal for a programmer interested in expeditious debugging. Nonetheless, much will be learned from the process of developing, implementing, and testing the scheduler. Even with limitations, a scheduled environment will represent a considerable improvement over a strictly space-shared environment. In this part of the project, Cray Research will develop the checkpoint and swapping software, and LLNL will develop a prototype scheduler along with the X-window interface that will allow a user to see which jobs are resident, which ones are waiting, and which processors are being used.

Materials Design	
Materials modeling	Will enable industry to simulate processes, such as casting and welding, that are involved in making new materials.
Advanced materials design	Will make available to industry advanced materials modeling codes. Codes will be ported, optimized, and integrated into Cray Research’s UniChem computational chemistry system.
Microwave components design	Will provide state-of-the-art microwave simulation exploiting the latest advances in computing capabilities.
Advanced Manufacturing	
Hydrocode library	Will address specific hydrodynamic simulation problems; make practical the analysis of 3D, reactive, multiphase flows common in U.S. industry; and provide adequate resolution of physical and chemical models.
Fluid–structure interaction	Will create a new time-domain simulation capability for structural acoustics by the shipbuilding and aerospace industries.
Fluid dynamics and structural response	Will allow the entire U.S. metals industry to develop improved forging, extrusion, and casting processes.
Shallow-junction devices	Will allow predictive modeling for the fabrication of next-generation semiconductor devices, reduce development time, and reduce cost.
Tools to Maximize the Use of MPP	
Dynamic time-sharing scheduler	Will create a more user-friendly environment for debugging and production.
MPP performance measurement	Will allow for the more efficient use of high-performance computers through better data collection, analysis, and visualization tools.
Portability tools	Will provide a means for writing portable applications, thus removing a major barrier to industrial use of MPP computers.

Workers at Los Alamos will address the need for portability tools. On conventional computers, users are accustomed to applications that can be moved to (ported) and executed on a variety of computers and workstations. In contrast, portability standards do not yet exist for massively parallel machines. Users must either execute their programs exclusively on one vendor's hardware, or they must maintain separate sources for separate machines. An inability to write portable code and the lack of portability standards and tools have been major barriers to the adoption of massively parallel machines by industry.

A broad objective of the ICI is to establish a common set of conventions for programming massively parallel machines, including common mathematical library calls, common message-passing interfaces, common parallel extensions, and so forth. The basic idea is to ensure that there is a common subset of languages and libraries that are supported on MPP machines. Portable applications will not only benefit the vendor and scientific communities, but they will also remove a major barrier to the industrial use of MPP computers.

Potential Results and Benefits

Benefits to U.S. Industry

The single, most important achievement of the project will be physically realistic simulations using advanced models that run routinely on massively parallel computers. Ultimately, such applications can be used by industry to help the U.S. to compete effectively in the global marketplace. The applications, concepts, and tools that will be transferred to industry will enhance competitiveness by reducing product cycle times, producing higher-quality products, requiring fewer prototypes, and developing more efficient industrial processes.

Another possible gain from the ICI is best understood from the perspective of an industrial partner in the project. All of the industrial partners will have invested significant staff time in developing certain applications. However, from their point of view, the ICI still represents a high-leverage situation. As the codes evolve, industrial participants will be able to judge for themselves whether or not the applications provide a competitive advantage. If they do, then the partners will have employed an optimal strategy to uncover what might otherwise have been an expensive question to answer.

With the considerable cost-shared contribution by Cray Research, the infrastructure for this project is provided by the laboratories at no cost to the other industrial partners. In addition, the laboratories are making available a constellation of computational experts and specialists in parallel computing. Therefore, it is possible for a company to develop a large-scale application and determine its usefulness without the major investment associated with building a local infrastructure—an investment involving tens of millions of dollars. With the development of a major code, a partner is free to decide if an investment in a dedicated machine for the purposes of production is beneficial. In effect, the expensive research and development costs and risks associated with such work will have been reduced to nearly zero.

Over the short term, our applications will address a few, critical industrial problems—ones also of interest to the laboratories—that have been only partially solvable to date. The box on pp. 10–11 summarizes the results we can expect from each application. The benefit of each one is substantial; taken as a whole, the benefits are as broad as the spectrum of individual topics. The payoffs range from increased cost

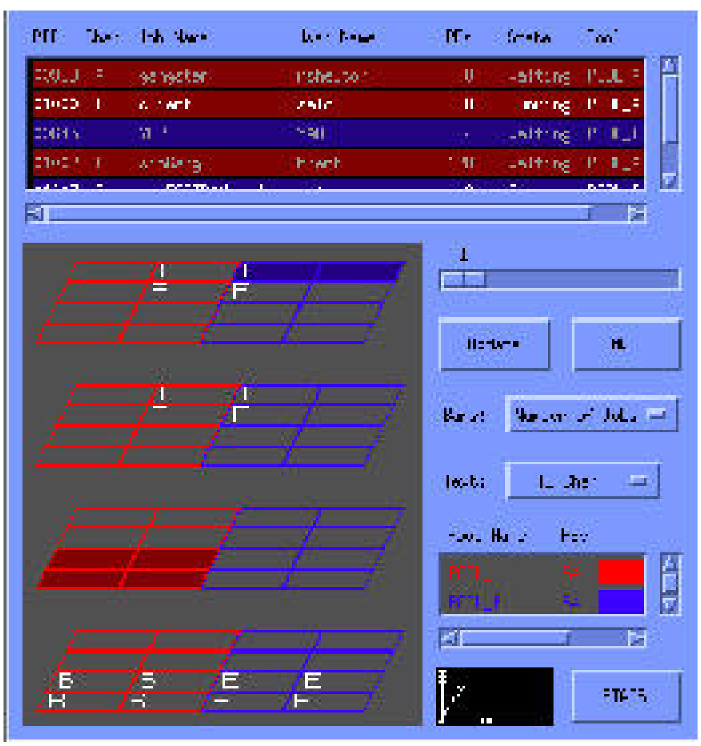


Figure 10. The Livermore Gang Scheduler will provide time-sharing support for massively parallel processor architectures. Such support greatly enhances the development environment on the most advanced systems and allows work to continue on production codes. In the display at the top, the status of each job and which processors are being used can be seen at a glance.

effectiveness in the design stage of a particular product or device to greater realism in three-dimensional (3D) models of chemical transport in soils or through the global atmosphere.

For example, many U.S. industries and the DOE face the enormously complex and expensive task of cleaning up contaminated soils and groundwater at thousands of facilities. Traditional two-dimensional models cannot adequately characterize the variations in subsurface properties that affect how contaminants migrate. In one of our research projects (described in more detail on pp. 24–25), we will develop a sophisticated 3D model to rapidly and accurately simulate fluid flow and chemical transport in heterogeneous, porous soils around waste sites.

Another effort in which LLNL and AT&T researchers will work together combines accurate simulations at the atomic scale with modern 3D visualization techniques to model diffusion mechanisms in silicon (Figure 11). Predictive modeling for the fabrication of next-generation semiconductor devices will reduce development time and cost. This work could have a significant effect on the U.S. semiconductor manufacturing industry.

Benefits to the DOE

The DOE's national laboratories have invested hundreds of millions of dollars in the computational software and hardware needed for their energy and defense missions. The phenomena that are modeled—often through techniques developed at the laboratories themselves—include materials properties, structures, pollution migration, charged-particle transport, shocks and detonation, fluid dynamics, acoustic and elastic wave propagation, combustion, plasma physics, nuclear physics, and electromagnetism.

The ICI serves a dual-use function. Applications originally created specifically for DOE programs will now be transformed into efficient massively parallel applications for use by U.S. industry. The modified and enhanced codes can be simultaneously reintegrated into ongoing DOE programs. For example, the parallelization of existing codes will serve the reliability and safety programs at LLNL. Furthermore, scientists not directly involved in the ICI will have access to the new techniques and machines for their own code development (see the box on p. 14). Such access and the acquisition of skills associated with new computational techniques can facilitate a wide range of ongoing DOE programs, including genetics research with its need for vast computational resources, environmental modeling, and weapons-related work.

To carry out its mission related to the Comprehensive Test Ban, the DOE requires more extensive use of advanced simulations. The ICI gives the DOE one more way to evaluate a possible path to massively parallel computing through cost sharing with computer vendors and other project participants. In addition, the kind of production environment we will develop can serve as a model for an MPP production environment on classified machines.

Benefits to All

A unique aspect of the ICI is that it will create an infrastructure that couples the most advanced MPP resources at the two national laboratories. The prototype interconnection by a high-speed network will allow researchers at either the Livermore or Los Alamos locations to access the resources of both sites using distributed computing techniques. Not only will the prototype environment increase the total computing capability at each location, but it will also give industry the benefit of experimenting

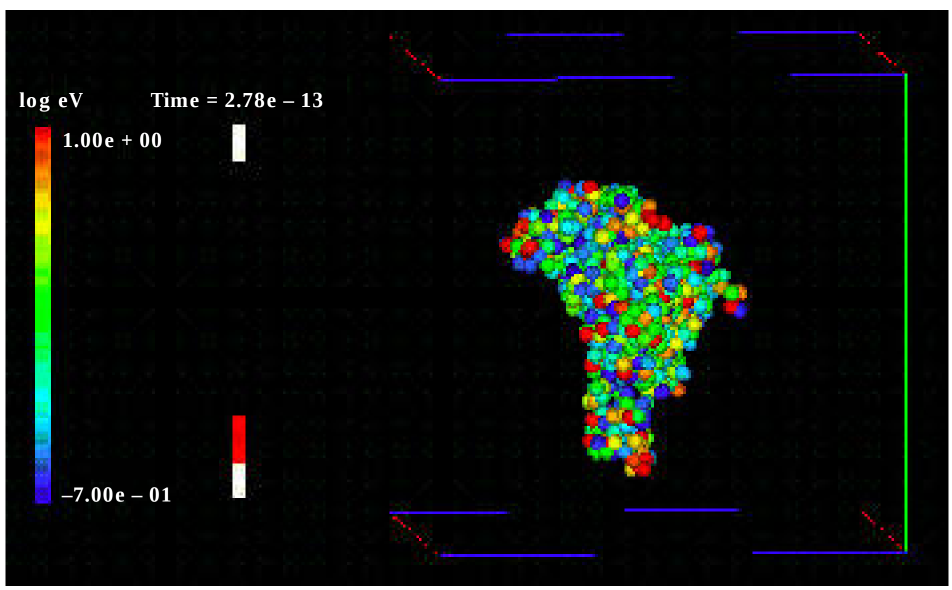


Figure 11. Predictive modeling for the fabrication of next-generation semiconductor devices will reduce development time and cost. Here, a silicon crystal is bombarded by an ion, and color is used to show the energy distribution. When doping crystals, this type of simulation helps us understand how atoms are displaced and how defects arise.

Building on the ICI

The machine resources at LLNL for the Industrial Computing Initiative (namely, the 128-processor CRAY T3D machine with 88 GB of disk) were chosen to accommodate nine different industry projects at LLNL. However, we will have little capacity to accommodate more ICI-like projects. To add worthy projects, such as the grand challenge problems in tokamak physics and climatology mentioned in this article, LLNL has decided, as an institutional project funded through the Director's Office, to double the number of processors and the disk capacity on the T3D. Sixteen processors will also be added to the 48-processor Meiko CS-2 parallel computer sited in the open environment.

The T3D enhancement will be provided in the same cost-shared manner as the first 128 processors, through a special arrangement with Cray Research. The resources available on two parallel computers will allow for ample expansion of ICI-like projects and for access by researchers at LLNL and within the UC system interested in developing high-performance massively parallel research projects. The environment across the two machines will be coordinated to allow, for instance, file sharing through a common file system as well as some common programming models. In this manner, researchers engaged in code development will be able to migrate and port across systems with minimal disruption. Some sharing of the tertiary storage environment at the Open Computing Facility is also planned. The Laboratory is particularly interested in using the enhanced computational resources to develop additional partnerships with industry and in developing collaborations with researchers at the UC campuses.

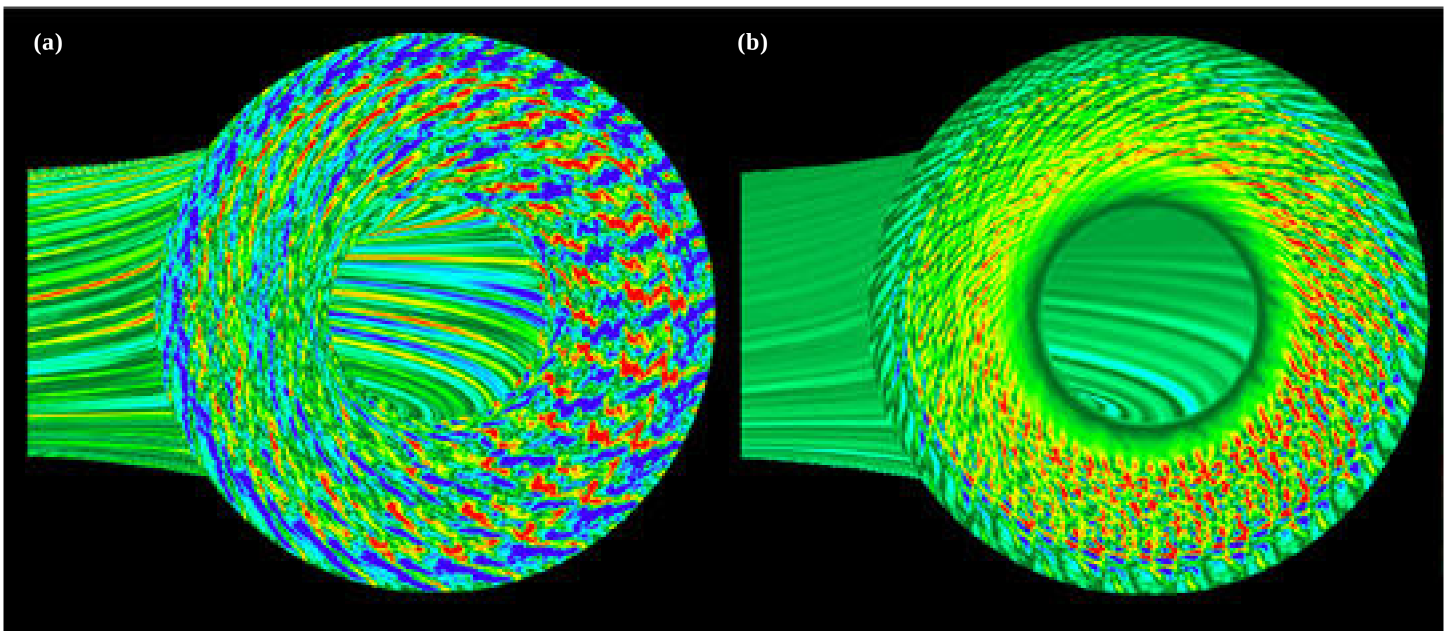


Figure 12. The Numerical Tokamak (NT) Project represents a grand challenge problem in the area of plasma physics. Along with other NT consortium participants, we are developing advanced computational models of tokamak physics using the most powerful high-performance computers. (a) This physical model shows how fusion ignition is prevented by the turbulent mixing of hot core plasma with cooler edge plasma. Red represents positive density perturbation; blue corresponds to negative perturbation. The view shows transient turbulent structures in cross section. (b) Numerical models explore the conditions that can reduce turbulence. Although this work is not part of the ICI, tokamak physics codes and those for other grand challenge problems will benefit from what is learned through the current ICI effort. (This calculation was done on a CM-5 machine at the Los Alamos ACL by LLNL researchers; results are displayed with a parallel ray-tracing technique developed here.)

with a distributed MPP computing environment before making the investment required to implement one. Coupling several dispersed machines to attack a single, but very large, application is of great interest to industry and the nation as a whole because it represents an economic approach to addressing large-scale computational needs.

Learning how to develop efficient parallel codes will also allow us to better address the problems identified as “grand challenges.” These problems are of such generality and complexity that they can be addressed only on large parallel computers; standard vector machines do not have sufficient speed and memory, and loosely coupled multicomputers do not have sufficient interprocessor communication efficiency.

An example of a grand challenge problem in the area of plasma physics is the Numerical Tokamak Project. In the 21st century, magnetic fusion could be the source of large amounts of electricity without contributing to global warming or acid rain. The U.S. goal is to have a magnetic fusion Demonstration Power Reactor on line by the year 2025. Now, a multi-institutional, multi-disciplinary collaboration, including researchers at LLNL, is developing some of the most advanced computational models of tokamak fusion devices to identify the most cost-effective design (Figure 12).

Another grand challenge problem involves global climate modeling. As part of a major international effort, LLNL investigators are trying to determine why different global climate models give dramatically different results even when they use the identical set of data. Grand challenge problems, such as the Numerical Tokamak Project, global climate modeling, and others, have great significance in addressing society’s problems and needs. Although these particular problems are not a part of the ICI, grand-challenge codes will benefit from what we learn through the ICI.

Summary

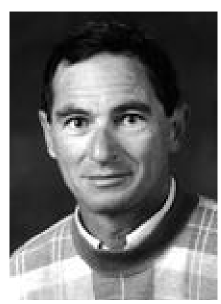
The widespread use of massively parallel computing as a scientific and industrial tool has been impeded by technical problems associated with available hardware and software. U.S. industries do not have the applications and necessary software tools and languages to facilitate use of MPP environments. To address this need, the Industrial Computing Initiative, which is part of the High-Performance Parallel Processing Project, is developing a set of efficient applications that can realistically simulate complex problems and are written in a way that allows them to run on MPP hardware from various vendors.

The ICI involves more than 40 scientists from Lawrence Livermore and Los Alamos national laboratories, six specialists in parallelization from Cray Research, two specialists from Thinking Machines Corporation, and nearly a score of industrial scientists. The delivery of a set of efficient parallel applications, serving as guideposts for subsequent work, can help U.S. industry compete more effectively in the global marketplace. The new applications, together with improved methods and tools, can reduce product cycle times, produce higher-quality products, and speed the development of more efficient industrial processes. At the same time, the new computational advances in massively parallel computing technology can be reintegrated into ongoing programs at the national laboratories to serve the missions defined by the DOE and to address national needs.

Key Words: computer performance; grand challenge problems; High-Performance Parallel Processing Project (H4P); Industrial Computing Initiative (ICI); massively parallel processing (MPP); parallel processing.

Notes and References

1. F. Baskett and J. L. Hennessey, “Microprocessors: From Desktops to Supercomputers,” *Science* **261**, 865 (1993).
2. This graph, developed by LLNL’s Eugene Brooks, is an extension of a projection first made by R. Turn (Ref. 3) 20 years ago and more recently used by Baskett and Hennessey (Ref. 1).
3. R. Turn, *Computers in the 1980s* (Columbia University Press, New York, N.Y., 1974), p. 80.
4. D. Bailey et al., *The NAS Parallel Benchmarks*, NASA Ames Research Center, Moffett Field, CA, RNR Technical Report RNR-94-007 (1994); D. H. Bailey et al., *NAS Parallel Benchmark Results 3-94*, NASA Ames Research Center, Moffett Field, CA, RNR Technical Report RNR-94-006 (1994).



For further information contact Michel G. McCoy (510) 422-4021.